



Development and Assessment of Upper Basic Students' Conceptual Understanding of Mathematics Knowledge Using Two-Tier Diagnostic Test in Kogi State, Nigeria

SUNDAY YOMI, SALAMI DHARRIYAT

Confluence University of Science and Technology, Osara, Kogi State, Nigeria

OLORUNNISHOLA OLUFIKAYO AGNES, SHUAIBU KENNEDY

Prince Abubakar Audu University, Anyigba, Kogi State, Nigeria

Abstract. Assessment of students' mathematics learning outcomes and instruments used in measuring the learning outcome are fundamental to the teaching and learning process of mathematics at the upper basic level of education. This study investigated development and assessment of upper basic students' conceptual understanding of mathematics knowledge using two-tier diagnostic test in Kogi State Nigeria. The study adopted instrumentation design, the population of the study were all the (50, 851 upper basic students of Kogi state. The sample consist 668 upper basic students which were selected using multistage sampling techniques. Two-Tier Mathematics Diagnostic Test (TTMDT) was developed by the researchers in three stages. the instrument TTMDT was trial tested on students outside the target population to establish the reliability of the instrument. Kuder-Richardson Formula 20 was used to measure the internal consistency reliability of the instrument. Data collected were analyzed quantitatively using frequency counts and percentages as well as descriptive statistics of mean and standard deviation to answer the research questions. The hypotheses were tested using t-test and analysis of variance (ANOVA) at 0.05 level of significance. The study adopted partial credit model in scoring. The findings of the study indicates that two-tier diagnostic test is valid and reliable to distinguish students who understand the concepts from those who misunderstood, and with the use of Rasch analysis two-tier test items could be improved upon by rewording items that showed anomalies base on the individual-item fit. It was recommended that teachers should adopt two-tier diagnostic test in assessing their students in mathematics.

Keywords: Development, Two-Tier, Conceptual Understanding, Upper Basic, Diagnostic Test.

1. Introduction

The role of Mathematics towards realizing the nation's scientific and technological aspiration is indisputable. According to Sunday and Haruna (2024) Mathematics is integral to daily life and everyone's long-term planning. Thus, Mathematics plays a significant role in human kind's ability to survive on Earth. This likely explains why arithmetic is taught as a required subject in elementary schools everywhere, including Nigeria. As such, no nation can thrive well without mathematics. This implies that Mathematics has much to offer in solving the problems of mankind by being the gateway to scientific and technological development of any nation. Consequently, the need to develop the necessary mathematical, scientific, and technological knowhow of the next generation demands assessing their conceptual understanding. However, for the purpose of this study, the upper basic mathematics curriculum will be considered, because at this level so many corrections can still be easily affected on a learner's mind.

Given the significance of mathematics, it is essential that students learn and understands Mathematics concepts. While concept is generally defined as an abstract or generic idea generalized from particular instances (Merriam-Webster's Dictionary, 2012), Ben-Hur (2016) notes that concepts are mental structures of intellectual relationships. Therefore, conception is the ability to form or develop correct mental representations of fundamental mathematical

ideas. This implies that conceptual understanding is knowledge rich in relationship and understanding; a connected web of knowledge; a network in which the linking relationships are as permanent as discrete bits of information and cannot be learned by rote, but thoughtful and reflective mental activity. Students with conceptual understanding isolate facts and methods. A student is not just taught how to do the mathematics but also the why behind doing it. Students are encouraged to see the bigger framework that underlies all fundamental mathematics ideals and to think fluidly so that they are able to apply their mathematics skills to a wide variety of problems. Ben-Hur (2016) maintains that conceptual understanding involves knowledge of concepts and recognizing their applications. Thus, conceptual understanding is deep and justifiable knowledge of mathematical ideas and concepts that aligns with those of mathematics experts and, that which reflects students' ability to reason in settings involving the careful explanation and or application of concepts. Onah (2019) attributes students' poor knowledge or conceptions of Mathematical knowledge to their prior experience exposed to such as perception, language, and cultural background, peer groups, belief among others, giving rise to misconceptions or alternative conceptions. The resulting misunderstanding or misconceptions if not challenged, become integrated into students' cognitive structures and interfere with subsequent learning. As a consequence, students will persistently experience difficulty in integrating any new information within their cognitive structures, resulting in an inappropriate understanding of new mathematical concepts. This could lead to poor learning outcome.

Gender differences remain a prime area of research in mathematics education with studies being carried out all over the world. On a similar note, Awofala (2017) found out that male students outperformed their female counterpart on average in mathematics, while Anjum (2016) and Ogannaya; Okafor; Abonyi and Ugama (2016) findings revealed that female students outperformed male counterpart in Science and Mathematics subjects. However, studies such as Alex and Mammen (2014) reported that gender differentials in Mathematics among male and female is converging, hence they perform similarly as a result of their thinking levels. This view was upheld by Timayi, Ibrahim and Sirajo (2016) who found no statistical difference between the performance of male and female students in Mathematics. According to Adeniyi and Kajuru (2016), even students (male or female) considered as higher achievers and achieving equally can also hold some unexpected mathematical misconceptions.

A two-tier test is a two-level multiple-choice question that diagnose students' misconceptions and determine students' conceptual understanding (Mutlu & Sesen, 2015). While the first tier of the two-tier test examines content knowledge; the second tier examines the reasons or supporting conception underlying such knowledge. In the two -tier diagnostic test, the items are specially designed to identify alternative conceptions or misunderstanding in a defined area. It is different from the regular multiple-choice tests in that the test items are usually constructed based on students' misconceptions that might be identified through open-ended questions unlike regular multiple-choice test items. Two-tier tests are generally superior to conventional multiple-choice tests because they provide researchers with an understanding of students reasoning behind their answers (Onah, 2019).

Some researchers have also developed instruments to help in this area (Ramli, Talib, Hassan & Manaf, 2020 ; Irawan & Wilujeng, 2020) but very few were seen done in mathematics and the instrument developed was not strong enough to effectively assess junior students' conceptual understanding and identification of misconceptions of mathematics knowledge, because many used concept mapping and simple multiple – choice test. The need to assess the conceptual understanding of learners and to identify any misconception they may hold in Mathematics necessitated this study. Hence, the study seeks to develop a reliable and valid instrument to assess the conceptual understanding and misconception of mathematics knowledge in upper basic secondary school using two-tier diagnostics test in Kogi State.

1.1 Statement of the Problem

It is no longer news that mathematics is the bedrock that provides the springboard for the growth of science and technology. Adequate use of Mathematics in modern science and technology is a sure way to harness all the natural resources (human and materials) which advances the economy of the nation. However, despite the attention mathematics gets, not much improvement in terms of students' performances has been recorded. For a long time, Mathematics is being seen as a difficult subject to learn, yet, it is needed for a meaningful and effective living in this modern age (Gafoor & Kurukkan, 2015). Shartz (2014) reported that students view mathematical concepts as the most difficult to learn. This has hindered many children from achieving the desired outcome in their mathematics performances. The (Basic Education Council Examination (BECE) results from 2018 to 2022 in Kogi State (Ministry of Education, Lokoja, 2023, This indicates that a lot is still needed to be done

for mathematics knowledge of upper basic secondary school students to come up to its desired height. As mathematics is a requirement of further study of the sciences, it has affected a large number of students in science related courses. Thereby, bringing set back in National development of the nation.

Some other researchers have also tried to develop instrument to cater for the proper assessment of the conceptual understanding of students but some of the instruments are limited to specific topics. It is against this background that the object of this research is to develop a valid and reliable instrument to assess upper basic school students' conceptual understanding in mathematics knowledge using two-tier diagnostics test in Kogi State.

1.2 Research Questions

The following research questions guided the study:

- What are estimates of item parameters of the Two-tier Mathematics diagnostic test (TTMDT) of upper basic student developed by the researchers?
- What is the internal consistency reliability estimate of the two-tier diagnostic test instrument developed to assess upper basic students' conceptual understanding as measured using a two-tier diagnostic test instrument?
- What percentage of upper basic students hold right conception as compared to those holding misconceptions of mathematics knowledge in Kogi State as measured using TTMDT?
- What is the mean difference between Male and Female upper basic students' conceptual understanding of Mathematics knowledge using anchor items in TTMDT?

1.3 Hypotheses

The following hypotheses were tested at 0.05 alpha level of significance

H₀₁. There is no significant difference in the scores of upper basic 1, 2 and 3 students' conceptual understanding of Mathematics knowledge using TTMDT.

H₀₂ There is no significant difference between upper basic 1 male and female students' conceptual understanding of Mathematics knowledge using TTMDT.

2. Literature Review

2.1 Theoretical Framework

2.1.1 Constructivism Theory by Jean Piaget's work in 20th Century.

Constructivism theory was founded by Jean Piaget (1896-1980). As a biologist, he was interested in how an organism adapt to the environment and how previous mental knowledge contributes to behaviour. It is a theory in education which posits that individuals do acquire knowledge and understanding by actively perceiving it within a direct process of knowledge transmission. Learners construct new understandings based on past knowledge through experiences. As events occur, each person reflects on their experiences and incorporate the new ideas with their prior knowledge. The root of constructivist learning theory date back to Jean Piaget's work in 20th century. He suggested that through processes of assimilation, human build their own knowledge. Jean Piaget is known as one of the first theorist in cognitive psychology. His theories indicate that humans create knowledge through the interaction between their experience and ideas.

This theory is very much related to this study in the sense that learners are encouraged to get conceptual understanding of whatever is meant to be done before it is been done. That is, the knowledge one has beforehand is instrumental to creating new improved knowledge, which also means that if a learner has misconceptions in any area of mathematics, it must be corrected before new knowledge can be created. This theory leads to changes in learners' schemas. And subsequently encourage conceptual understanding of the concepts in mathematics.

2.1.2 Item Responses theory

Item response theory (IRT) has its roots in Thurstone's work to scale tests of "mental development" in the 1920s. Though in several other people's works, its origin can be traced back to the early 20th century when pioneers like George Rasch and Frederic Lord made significant contributions to its development. Rasch introduced the concept of item Characteristic Curve (ICC), which forms the basis of IRT models. Thurstone's method is based on the assumption of a latent scale underlying the responses on which both the items and individuals can be mapped. No doubt his idea of scale was inspired by earlier developments in psychophysics in which probabilities of psychological response to physical stimuli are modeled as a function of the strength of the stimulus. The foundation of IRT

is a mathematical model defined by item parameters. For dichotomous items (those scored correct/incorrect), each item has three parameters:

a: The discrimination parameter, an index of how well the item differentiates low from top examinees; typically ranges from 0 to 2, where higher is better, though not many items are above 1.0.

b: The difficulty parameter, an index of what level of examinees for which the item is appropriate; typically ranges from -3 to +3, with 0 being an average examinee level.

c: The pseudo-guessing parameter, which is a lower asymptote; typically, is focused on $1/k$ where k is the number of options.

In addition to being used to evaluate each item individually, IRFs are combined in various ways to evaluate the overall test or form. The two most important approaches are the conditional standard error of measurement (CSEM) and the test information function (TIF). The test information function is higher where the test is providing more measurement information about examinees; if relatively low in a certain range of examinee ability, those examinees are not being measured accurately. The CSEM is the inverse of the TIF, and has the interpretable advantage of being usable for confidence intervals; a person's score plus or minus 1.96 times the SEM is a 95% confidence interval for their score.

This present study has developed an instrument and uses a branch of Item Response model (Rasch) to check the psychometric analysis of the instrument and at the same time use it in analyzing reliability of the instrument. This theory is relevant to this study because it involves creating or developing an instrument that overcome classical testing limitations.

2.2 Upper Basic Secondary Schools Mathematics Curriculum in Nigeria

Reformation in school mathematics curricula goes along side with the reforms in Nigeria education system. Reform in education, would mean attempting to evolve a system of education that will respond positively to the needs and demands of the changing nature of the society. The growth of knowledge and techniques required for the industry, science, commerce, medicine and technology continue. All the future needs cannot be given or satisfied today but a solid foundation can be laid in readiness for the demands of these disciplines. To conform to these growing social needs, mathematics has been scrutinized and found to need changes. Thus, the so-called traditional mathematics (Arithmetical process) become grossly inadequate to cope with the modern needs, which gave room for the introduction of

modern mathematics. In the first half of the 20th century, curriculum development emphasized shop - and - yard skills prompted by the idea of functionalism (education you can see), some educators focused on identifying minimal competencies needed to perform different jobs; dollars - and - naira mathematics for clerking, feet - and - inches mathematics for carpentry, measuring cup - and- spoons mathematics for cooks and home mathematics new 9-3-4 system of education in Nigeria which aims at reinforcing the extant 6-3-3-4 system of education introduced in 1983, consist of the first nine years of basic and compulsory education up to the Jss3 levels, three years in the senior secondary school and four years in the tertiary institutions. Its first 9- years duration which is popularly known as Universal Basic Education was designed to meet the millennium development Goals (MDGS) and Education for All (EFA) by 2015 (Federal Republic of Nigeria, 2015).

2.3 Conceptual Understanding

Mathematics curriculum is made up of concepts, skills and generalizations. A concept is an abstract or generic idea. It is a symbolic representation (almost always verbal) used in the process of abstract thinking and possessing a general significance, corresponding to an ensemble of concrete representations with regard to what they have in common. This implies that a concept is characterized by its expression of an idea, a general ideal representation of a class of objects, based on their common features. Thus, concepts are mental models that may be constructed by the learners. Each of these concepts, according to Nisa, Waluya, Kartono and Mariani (2020) is an abstraction, namely a mental image of an object received (for example a triangle), a mental process that becomes a concept (such as counting into numbers), and a formal system (such as a permutation group). Which is based on its properties with a concept built through deductive logic. A mental model of concepts constructed by learners which are inconsistent with the experts' opinion, result in a misconception held by the learners. Meanwhile, a perfect grasp of concept brings about the term conceptual understanding. There is a world of difference between a student who can summon a mnemonic device to expand a product such as $(a + b)(x + y)$ and a student who can explain where the mnemonic comes from. Mill (2016) succinctly points out that one of the conditions necessary for conceptual understanding to take place is that misconceptions if any, must be corrected. Therefore, this study is carried out to determine area of junior secondary school student's conceptual understanding and misconception in mathematics knowledge using two-tier diagnostic test in Kogi State.

2.4 Open-ended Tests

In order to investigate students' understanding, open-ended free-response tests are commonly used. In open-ended tests, students are given the opportunity to state in writing everything they know about the problem/subject/concept. Students are also asked to state the reason/justification for their answers to the problem. This allows the teacher to identify possible misconceptions. Distractors in multiple-choice tests are prepared on the basis of students' responses to test question and to other open-ended questions. This method gives students more time to think and write about their idea. It is to identify what students know and why they believe it to be valid, but it is difficult to interpret and analyze the results of the open-ended questionnaires (Gurel, Eryilmaz & McDermott, 2015). Furthermore, because of language problems, identification of students' misconceptions becomes difficult and more so that students are less eager to write their answers in full sentences (Gurel & Eryilmaz, 2015).

2.5 Multiple-choice Tests

Multiple-choice tests have been used for measuring students' understanding of concepts as they enable a large number of students to be sampled in a given amount of time. It also helps to overcome the difficulties encountered in open-ended testing processes. These tests are also easy to administer and accurate in scoring and the results obtained are also easily processed and analyzed (Gurel, Eryilmaz & McDermott, 2015). A typical multiple-choice test item has a stem, which may be a statement or a question and four or five possible answers of which only one is correct. A student can give a correct answer with a wrong reason or a wrong answer with a correct reason (Kirbulut & Geban, 2014). Therefore, to overcome the limitations of multiple-choice tests in assessing conceptual understanding and identification of misconceptions among students, researchers have proposed the use of multiple-choice tests items with distractors, which are used based on students' answers to open-ended questions and/or other essay questions (Kanli, 2014). Kanli (2014) found the use of justifications when answering multiple-choice test items to be a sensitive and effective way of assessing meaningful learning among students, and addresses the extent of the limitations of traditional multiple-choice test items. The positive outcomes of findings related to students' justifications to test items led to the development of the two-tier multiple-choice diagnostic instrument, specifically for the purpose of measuring conceptual understanding and identifying students' misconceptions in limited and content areas.

2.6 Two-tier Diagnostic Tests

The two-tier diagnostic instrument is a multilevel assessment model in knowing the weakness and mistakes of students in understanding a concept. Two-tier diagnostic tests consist of two sections at least. The first section aims to diagnose how the individual interprets discipline knowledge. In the second section, students are asked to state the reason(s) for their answer in the first section (Kanli, 2015). The first tier of each multiple-choice item consists of a content question having usually two to four choices, one of which is the correct answer and the rest are distractors. The second tier of each item contains a set of usually four possible reasons for the answer given to the first part. The reasons consist of the designated correct answer, together with identified students' conceptions and/or misconceptions. When more than one alternative conception is given, these are included as separate alternative reason responses. The use of two-tier test is aimed at reducing the guessing factor of students because students are required to provide reasons to the answers they choose. (Lengkong, Istiyono., Rampean, Rejeki, Tumanggor&Nirmala,2020).

2.7 Empirical Studies on Conceptual Understanding of Mathematics

Yang, Leung, and Zhang (2019) investigated junior secondary school students' conceptions understanding and approaches to learning mathematics and the relationships in Mainland China. The study employed quantitative and qualitative research design. Two factors of students' lower-level conceptions of learning mathematics, "memorizing" and "testing", were the strongest predictor for the surface approaches to mathematics learning, while students' higher-level conceptions of learning mathematics, such as "applying" and "understanding and mathematical thinking", had a noticeable effect on their deep approaches to learning mathematics. However, under the pressure of examination in Mainland China, "understanding and mathematical thinking" was also found to exert quite a strong influence on students' "surface motive".

Malatjie and Machaba (2019) explored learners' conceptual understanding of coordinates and transformation geometry through concept mapping in South African University. A qualitative case study design was employed in this study, data were collected using an investigative task, observation and reflective interviews on a sample of 34 Grade 12 Mathematics learners. The study findings revealed that, although

some learners struggled with linking words and omitted some concepts in their concept maps, there were some indications of conceptual understanding of coordinate and transformation geometry when the learners were probed during the interview. Thus, the learners' conceptual understanding of coordinate and transformation geometry could be improved when they are taught, using concept mapping.

Empirical Studies on development and usage of Two-Tier Diagnostics Test for accessing conceptual understanding and misconception

Several studies have been done in connection with students' conceptual understanding using a two-tier diagnostic instrument.

Aliyu, (2015) developed and validated Mathematics Achievement Test using the Rasch model. An instrumentation research design was adopted. The result showed that 65 items not only met the Rasch model assumption of measurement construct (fitting and invariant) but also demonstrated good psychometric properties. The result showed that items and person separations indices were 13.17 and 2.93 while item and person reliability were 0.99 and 0.78 respectively. The MNSQ for both infit and outfit were 0.94 and 1.08 respectively while the ZSTD for both infit and outfit are -1.7 and +2.0 respectively which were within the acceptable range of 0.7-1.1 for MNSQ for sample > 1000 while -2.0 to +2.0 for ZSTD. The difficulty level of the items ranges between -1.95logit to 7.45logit. Also, the output results were expressed in both wit and logit units.

Aligba and Iorja (2021) Assessed conceptual and procedural knowledge of students with special needs in Mathematics in Benue State Algebraic Diagnostic Test (TTADT) item cycle I and cycle II were adapted and validated. The reliability of TTADT was calculated using Pearson Product Moment Correlation Coefficient and was found to be 0.96. Mean and Standard Deviation were used to analysed data to answer the research questions, while t-test statistics was used to test the null hypotheses at 0.05 level of significance. The finding of the study shows that the Mean scores of the SSN in Concept Knowledge and Underlining Reasoning were very low below 40% (31.25% and 21.08% respectively), while that of Procedural Knowledge is 40.69% and that there exist a significant difference between the performance of SSN in Conceptual Knowledge and Underlining Reasoning in Algebra in favour of Conceptual Knowledge ($t=5.71$; $P=0.00<0.05$).

Lengkong, Istiyono, Rampean, Rejeki, Tumanggor and Nirmala (2020) in their research to develop a two-tier test instrument to diagnose students' conceptual understanding abilities and knowing the level of student misconceptions on the topic of simple harmonic motion. The test instrument developed and tested to 60 students of Class X MIPA was able to measure the error of students by 30.5% having misconceptions and by 27.6% not understanding the concept. Of the nine items given with different conception, the largest percentage of students' misconceptions is in item number 4, which is 68.3% of students experiencing a misconception about the relationship of spring length to their frequency value.

Thus, it was clear that from the above reviews that the use of two-tier diagnostic test is effective in assessing the conceptual knowledge and in identifying students' misconceptions about knowledge. Only few of these studies above took place in Nigeria but in other part of Nigeria and mostly all in pure sciences. Therefore, this study seeks to find out whether the two-tier diagnostic test that was developed by the researcher was reliable and valid to access the students' conceptual understanding and identify the misconceptions in Mathematics in Kogi State, Nigeria.

3. Research Methodology

3.1 Research Design

The study adopts instrumentation design. It is a design which typically involves creating tools to measure or assess something. For the purpose of this study, instrumentation test survey design was employed. It is a type of research study that focuses on introduction of new or modified content, procedure, technologies or instruments of educational practice (Ihekwaba, Nkwocha & Unamba, 2019)..

3.2 Population

The population of this study consists of all students in Upper Basic classes of the secondary Schools in the twenty-one (21) Local Government Area of Kogi Stat, Nigeria. Their total number stands at fifty thousand, eight hundred and fifty-one (50, 851) students for the 2024/2025 academic session. There are 26,184 males and 24,667 female students

3.3 Sample and Sampling Technique

The study employed multi-stage sampling procedures to select six hundred and sixty- eight (668) students of the upper basic students of kogi State. The most used sampling method in this study under the multi-stage

procedure is the simple random sampling method. Simple random sampling is used to make statistical inferences about population (Thomas, 2023). It provides each individual or members of a population with an equal and fair probability of being chosen.

3.4 Instrumentation

Two-Tier Mathematics Diagnostic Test (TTMDT) was developed by the researcher in three stages adopting procedures by Treagust in (widiyatmoko & shimizu, 2018) was used for the study. The following procedures were involved; defining the content area of the study, obtaining information about students' conceptions by having students provide free response answers and reasons for their answers, which is the first cycle and several steps in developing the diagnostic instrument and the validation of final version which is the second cycle,

3.5 Defining Content Area of the Study

This stage involved the researcher defining the content boundaries, which will further be the knowledge required to understand the mathematics knowledge as specified in the Nigerian upper basic secondary school Mathematics curriculum.

3.6 Obtaining Information about Students' Conceptual knowledge (Open-ended Test)

An open-ended test items was first developed by the researcher as a Two-Tier Mathematics Diagnostic Test (TTMDT) cycle 1 to elicit students' conceptual knowledge from their written responses. The test items consist of only a stem requiring students to supply the answers and advance reasons for their answers. The test had twenty items for upper basic 1, twenty items for upper basic 2 and twenty items for upper basic 3. The purpose of the cycle 1 test was to identify students' conceptual understanding or misconception (correct and incorrect) which could serve as distractors for the construction of cycle 2 items. The test is designed to elicit students' spontaneous conceptions rooted deeply in the minds of the students the cycle two (2) is a multiple-choice instrument that has both the multiple-choice options to choose from and the underlying reasoning parts. To ensure construct validity, TTMDT was developed based on the table of specification

3.7 Development of the Two-Tier Multiple-choice Diagnostic Instrument

This was developed based on the conceptual understanding and misconceptions (responses)

obtained from the administration of cycle 1 test items. Each item in the open-ended test items on students was repeated to form a Two-Tier Mathematics Diagnostic Test (TTMDT) cycle II. Each test item in the instrument consists of a stem followed by a two-tier multiple-choice question. The first tier is the conventional content questions with four possible answers (one correct and three distractors) related to students' misconceptions in addition to the mathematically correct answer. The second-tier question consist of multiple-choice set of reasons students had given that is associated with the answers that were given to the first-tier in addition to the mathematically accepted reasons. Students' answer to each item will be scored correct when both the correct choice (answer) and reasons are correct with full credit (2 marks). Getting part only one correct earned a partial credit (1mark). The instrument consists of 20 two- tier multiple choice objective test items each for upper basic 1, 2 and 3 respectively. There are five (5) questions called anchor questions, items common to the three upper basic classes. Items were picked from topics from upper basic 1 and repeated in upper basic 2 and 3. With this, the researcher was not bias in verifying students conceptual understanding and the misconceptions across the three classes (See Appendix H pp. 180). Furthermore, the test items were subjected to psychometric analysis. This was an attempt to determine the quality of the test in terms of how difficult the test items may be using the Rasch approach. These were calculated by computing the item difficulty for each class of test items of (TTMDT) administered on students. Metha and Mokhasi (2014) recommend Item difficulty (I.D) 0.30 – 0.7 as acceptable range (See appendix I pp. 181) 20 items were submitted for the psychometric analysis and the 20 items per class were considered for the study. Since the distractors are based on students' responses from the administration of the open-ended questionnaire, the psychometric indices of distractors are excluded from the analysis.

3.8 Validation of the Instrument

The instruments TTMDT cycles I and II for the study were subjected to face, content and construct validation by five experts. Three (3) in Science and Mathematics Education Department, Benue State University, Makurdi and two experts in measurement and evaluation departments, from Prince Abubakar Audu University, Ayingba, Kogi State. The experts' advice was sought in terms of scope of coverage, content relevance, language level, ambiguity and vagueness of expression as well as suitability of items for upper basic mathematics topics specified. The advice and comments from the experts were used to

modify the items of the first cycle and the development of the final version of the items of the two-tier diagnostic instrument. The final version of the two-tier multiple-choice diagnostic instrument then emerged (cycle 1 and 2) and was used in this study. The modification instructions were attached.

3.9 Reliability

Reliability is the degree to which the result of a measurement, calculations, or specification can be depended on to be accurate. Research instrument is highly reliable if the test consistently measures what it is expected to measure under identical conditions. According to Aljaghsi, Abu, Algutaibi and Mundt (2021) the stability of scores over time requires all other things to be equal. That is, for test or instrument to have good reliability it means that the respondent will obtain the same score on repeated testing as long as no other extraneous affect the score.

The final version of the instrument TTMDT was trial tested on students outside the target population to establish the reliability of the instrument. Kuder-Richardson Formula 20 was used to measure the internal consistency reliability of the test in which each question only has two answers right and wrong. The value for KR-20 ranges from 0 to 1, with higher reliability. The students were tested and the scores gotten were used in establishing the internal consistency reliability coefficient of the instrument based on Kuder and Richardson formula (KR- 20) and the reliability coefficient stands at 0.78, 0.83 and 0.77, for upper basic 3, 2 and 1 respectively (See appendix K pp. 190). However, the reliability coefficient bases on IRT is also given as 0.76, 0.74 and 0.68 respectively.

3.10 Method of Data Collection

In collecting data, the researcher obtained a written permission from the university that was presented at the sampled schools for proper cooperation of the schools during the data collection exercise. Eighteen (18) Mathematics teachers were trained as research assistant during the exercise, three (3) Mathematics teachers per school. The Two-Tier mathematics Diagnostic Test (TTMDT) cycle II that was developed by the researcher was used for the data collection. The researcher organized a day orientation programme for the administration of the test items in each school. The programme Manual covered:

- The concepts of Two-Tier multiple-choice questions
- The essence of underlying reasons as the second - tier multiple-choice test items

- The use of TTMDT and the general conduct of the test administration

The programme took two (2) hours per day in each chosen school. One hour for the orientation and one hour for the administration of the instrument. These research assistants (three in each sampled school) assisted in the administration of the instrument and collection of the data. This reduced the cases of misuse of the instrument by the respondents. The instruments were administered on the students and the data collected immediately.

3.11 Method of Data Analysis

Data collected were analyzed quantitatively using frequency counts and percentages as well as descriptive statistics of mean and standard deviation to answer the research questions. The hypotheses were tested using t-test and analysis of variance (ANOVA) at 0.05 level of significance. The study adopted partial credit model in scoring. In this model, student received a score of 2 marks if he/she responded correctly to the first tier (content choice) and correctly to the second tier (reasoning part), signifying full conceptual understanding of mathematical knowledge. Also correct answer to the content part and wrong reason attracted a score of 1 mark, signifying partial understanding (misconception) while wrong answer and correct reason or wrong answer and wrong reason earned 0 marks, signifying no conceptual understanding or lack of knowledge This is because if students' ability is assumed to progress from knowing to explanation, students whose answer to first tier is incorrect but correct explanation should not exist on the progression path and this pattern can be attributed to guessing. Reliability analysis using Rasch model which relies on Item Response Theory (IRT) were used to answer the research questions 1 and 2 Rasch analysis is a paradigm for designing and testing instrument used for assessing abilities and other attributes and addresses measurement issues required for validating an outcome measure (like a test), including internal construct validity. In Rasch partial credit model, a reliability of at least 0.50 is required for a separation index of 1 (Aliyu, 2015).

4. Analysis and Interpretation

The data collected were presented and analyzed using Dichotomous Rasch Model software programme and each variable was coded as 0 or 1 with the type of numeric-continuous in jamovi, to answer the research question one and two. Frequency count and simple percentages were used to answer the research questions three, while means and standard deviations

were used to answer research questions four, five, six, seven, eight, nine, ten and eleven. T-test and ANOVA were used to test the hypotheses at 0.05 level of significance.

Research Question 1: What is the estimate of item parameter of Two-tier Mathematics Diagnostic Test (TTMDT) of upper basic 1, 2 and 3 respectively?

This research question was answered by using dichotomous Rasch Model software program to establish the psychometric property of the instrument. The results are presented in Table 1 and Fig 1 for upper basic 3, Table 2 and Fig 2 for upper basic 2 and Table 3 and Fig 3 for upper basic 1.

Table 1: Estimates of Item Parameter of Two-tier Mathematics Diagnostic Test (TTMDT) for Upper Basic Students

Table 1.0: Item Statistics

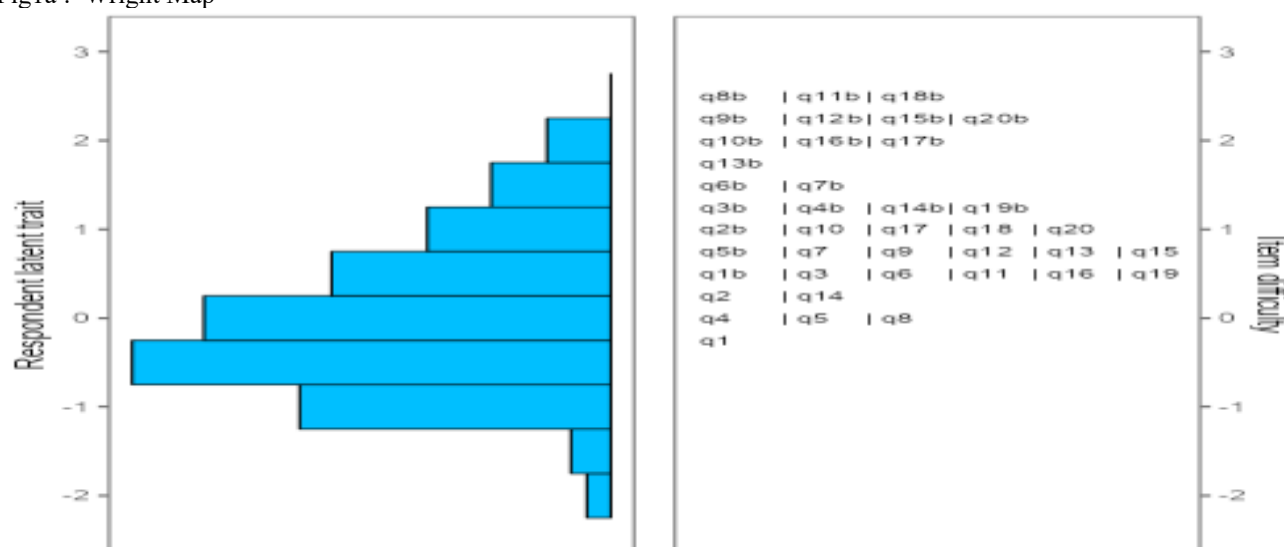
	Proportion	Measure	S.E.Measure	Infit	Outfit
q1	0.5314	-0.1486	0.137	1.054	1.105
q1b	0.3849	0.5192	0.141	1.111	1.143
q2	0.4603	0.1715	0.138	1.003	1.003
q2b	0.3054	0.9147	0.148	0.973	0.989
q3	0.3724	0.5791	0.142	0.979	0.980
q3b	0.2385	1.2921	0.160	0.946	0.927
q4	0.5021	-0.0170	0.137	0.973	0.956
q4b	0.2427	1.2667	0.159	0.938	0.918
q5	0.5230	-0.1110	0.137	1.013	1.005
q5b	0.3515	0.6808	0.143	0.910	0.888
q6	0.3891	0.4994	0.141	1.020	1.051
q6b	0.2050	1.5063	0.168	0.918	0.875
q7	0.3556	0.6603	0.143	0.989	0.986
q7b	0.2008	1.5347	0.169	0.972	0.975
q8	0.4728	0.1148	0.137	1.016	1.022
q8b	0.1004	2.4111	0.222	1.053	1.206
q9	0.3264	0.8064	0.146	0.983	1.049
q9b	0.1213	2.1841	0.205	0.954	0.884
q10	0.3013	0.9368	0.149	0.994	0.976
q10b	0.1339	2.0629	0.197	0.967	0.822
q11	0.4142	0.3819	0.139	0.981	0.976
q11b	0.0921	2.5132	0.230	0.981	0.885
q12	0.3222	0.8278	0.146	1.046	1.035
q12b	0.1130	2.2707	0.211	0.979	0.881
q13	0.3515	0.6808	0.143	0.953	0.928
q13b	0.1757	1.7146	0.177	0.949	0.917
q14	0.4435	0.2474	0.138	0.981	0.982
q14b	0.2469	1.2417	0.158	0.991	1.007
q15	0.3556	0.6603	0.143	1.008	1.004
q15b	0.1213	2.1841	0.205	1.001	1.044
q16	0.4142	0.3819	0.139	0.995	1.006
q16b	0.1297	2.1023	0.200	0.947	0.879
q17	0.2887	1.0041	0.151	1.133	1.150
q17b	0.1381	2.0246	0.195	1.132	1.274

Table 1.0: Item Statistics

	Proportion	Measure	S.E.Measure	Infit	Outfit
q18	0.2971	0.9590	0.150	1.003	1.043
q18b	0.0837	2.6236	0.240	1.054	1.148
q19	0.4142	0.3819	0.139	1.033	1.031
q19b	0.2301	1.3437	0.162	1.029	1.021
q20	0.3054	0.9147	0.148	1.028	1.072
q20b	0.1130	2.2707	0.211	0.908	0.870

Note. Infit= Information-weighted mean square statistic; Outfit= Outlier-sensitive means square statistic.

Fig1a : Wright Map



Data in Table 1 shows the Rasch item statistics where the first column identifies the items (q1, q2, q3..... q20) and reasons to each corresponding item (q1b, q2b, q3b.....qb20). Column 2 is the proportion, indicating the percentage of respondents that got the items (option and reason) correctly and this range from 0.11 to 0.53. The third column is the measure of item difficulty level, followed by the standard error (SE). The last two columns are the Infit and Outfit measures.

From the table, it shows that item 1 is the most correctly scored item with 53.14% to while item 17 was least correctly scored with 28.87% for the part a. The analysis showed that respondents find it difficult to provide correct reasons despite that the first items are correct. The range for correct reasons is from 0.08 to 0.38. Item 18 was the most difficult item for respondents to provide reason for correct option with only 8.37% correctly and the highest percentage (38.49%) of respondents got the correct reason on item 1. The third column (measure), further confirmed the easiest items (q4, q5, q1) with (-0.02, -0.11, -0.14) and most difficult items (q17, q18, q10, q20) with (1.00, 0.96, 0.94, 0.91). The acceptable threshold of Infit and Outfit (stable performance) ranges from 0.6 and 1.2, the spread of the values under Infit and Outfit are all within this range, which shows that all the items are pretty good, (Aliyu, 2015). Fig1 shows the wright map which is also known as item map. It shows the spread of item difficulty and person's ability. The item difficulty ranges from -0.02 to 1.00 for the options while the difficulty level of reasons for the correct options are 1.2 to 2.5. The figure also shows the person ability. The model fit indicates reliability coefficient of 0.76 which shows that items are good and very reliable

Research Question 2: What is the internal consistency reliability estimate of the two-tier diagnostic test instrument developed for assessing upper basic students' conceptual understanding and misconception of Mathematics knowledge?

Also using Dichotomous Rasch Model software program and each variable was coded as 0 or 1 with the type of numeric-continuous in jamovi, the research question two was answered.

Table 2: Person Reliability

Upper basic 3	Upper basic 2	Upper basic 1
0.76	0.74	0.68

Table 2 shows the model fit which indicate the person reliability. The person reliability of 0.76, 0.74 and 0.68 representing upper basic 3, 2 and 1 classes respectively shows that the items are good and reliable. K-20 gave a higher coefficient 0.78, 0.83 and 0.77, however the Rasch model software gave the most reliable ones because this software considered the all items, one after the other.

Research Question 3: What is the percentage of upper basic 1, 2 and 3 students holding right conception and those holding misconceptions of mathematics knowledge in Kogi State as assessed using TTMDT?

Table 3: Frequency and Percentage of Upper Basic 1, 2 and 3 Students Holding Right Conceptions of each Mathematical Concept in the TTMDT (n = 668)

	Items					
	Upper Basic 1 Tier 1	Tier 1 & 2	Upper Basic 2 Tier 1	Tier 1 & 2	Upper Basic 3 Tier 1	Tier 1 & 2
	No (%)	No (%)	No (%)	No (%)	No (%)	No (%)
1	58(36.2)	48(30.0)	124(46.1)	59(33.1)	127(53.1)	92(38.5)
2	92(57.5)	31(19.4)	182(67.7)	120(44.6)	110(46.0)	73(30.5)
3	84(52.5)	47(29.4)	133(49.4)	97(36.1)	89(37.2)	57(23.8)
4	72(45.0)	32(20.0)	170(63.2)	145(53.9)	120(50.2)	58(24.3)
5	63(39.4)	38(23.8)	125(46.5)	85(31.6)	120(52.3)	84(35.1)
6	90(56.2)	68(42.5)	140(52.0)	100(37.2)	93(38.9)	49(20.5)
7	77(48.1)	50(31.2)	130(48.3)	64(23.8)	185(36.6)	48(20.1)
8	85(53.1)	53(33.1)	191(71.0)	95(35.3)	113(47.3)	24(10.0)
9	70(43.8)	46(28.8)	97(36.1)	35(13.0)	78(32.6)	29(12.1)
10	91(56.9)	62(38.8)	87(32.3)	51(19.0)	72(30.1)	32(13.4)
11	83(51.9)	60(37.5)	126(46.8)	62(23.0)	99(44.4)	22(9.2)
12	66(41.2)	45(28.8)	112(41.6)	68(25.4)	77(32.2)	77(11.3)
13	56(35.8)	39(24.4)	106(39.4)	69(25.7)	84(31.5)	42(17.6)
14	68(42.5)	32(20.0)	131(48.7)	72(26.8)	106(44.4)	59(24.7)
15	70(43.8)	51(31.9)	126(46.8)	67(24.9)	85(35.6)	29(12.1)
16	63(39.4)	37(23.1)	177(65.8)	132(49.9)	99(41.4)	31(13.1)
17	43(26.9)	22(13.8)	235(50.2)	119(44.2)	69(29.9)	30(13.8)
18	69(43.1)	48(30.0)	151(56.1)	104(38.7)	71(29.7)	20(8.4)
19	68(42.5)	35(29.9)	136(50.6)	89(33.1)	99(41.4)	55(23.0)
20	39(24.4)	26(16.2)	135(50.9)	90(33.5)	73(30.5)	27(11.3)

The Results summary of the analysis in Table 3 revealed that 19 out of the 20 items for Upper Basic 1 have less than 40 percent (%) of the students with right conception. It also reveals that 16 out of the 20 items for Upper Basic 2 have less than 40 percent (%) of the students with right conception and finally the data revealed that all items for Upper Basic 3 have less than 40 percent (%) of students with right conception. This suggests high level of misconception in mathematics knowledge among Upper Basic students. The analysis also indicates that only 1 student of Upper Basic 1, 4 students of Upper Basic 2 and none of Upper Basic 3 students respectively had at least 40 percent of the items having right conceptions.

Research Question 4: What is the difference in the mean scores of Male and Female students' conceptual understanding using the anchor questions in TTMDT?

Table 4: Mean and Standard Deviation of Upper Basic Male and Female Student's Conceptual Understanding of Mathematics Knowledge Assessed Using Two-Tier Test

Sex	N	Mean (\bar{x})	Standard Deviation (SD)
Male	351	3.35	2.12
Female	317	3.69	2.22
Mean Difference		0.34	

The summary result of the analysis in Table 4 shows that the mean conceptual understanding score of male students is 3.35 with standard deviation of 2.12, while that of the female students' conceptual understanding score is 3.69 with standard deviation of 2.22. The result of the analysis indicates that the female students seem to edge the male counterpart in their understanding as their mean difference (0.34) is in favor of the female students.

The following hypotheses were tested at 0.05 alpha level of significance

Hypothesis 1

H0₁ There is no significant difference among the scores of upper basic 1, 2 and 3 students' conceptual understanding of Mathematics knowledge.

Table 5: One Way-ANOVA for Comparison of the Mean of Upper Basic 1, 2 and 3 in their Conceptual Understanding of Mathematics knowledge

understanding ²	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	62.602	2	31.301	6.727	.001
Within Groups	3094.302	666	4.653		
Total	3156.904	668			

The summary result of the analysis on Table 5 presents one-way ANOVA comparing the mean score of upper basic 1, 2 and 3 students on their conceptual understanding in mathematics. The result indicated that observe mean differences among the groups is considered significantly at $f = 0.304$, $P = 0.001 < 0.05$. Hence the hypothesis that there is no significant difference in the scores among upper basic 1, 2 and 3 students' conceptual understanding of Mathematics knowledge was rejected. This implies that there is significant difference in the score of upper basic 1, 2 and 3 students as assessed using TTMDT.

Hypothesis 2: There is no significant difference between male and female upper basic 1 students' conceptual understanding of Mathematics knowledge using two-tier diagnostic test

Table 6: T-Test Result of Male and Female Upper Basic 1 Students' Conceptual Understanding Scores of Mathematics Knowledge as Assessed Using TTMDT

Gender	No	(\bar{x})	SD	DF	t	Sig.	Discussion at $P < 0.05$
Male	74	14.50	5.93				
Female	86	14.23	5.54				
	158	.313	0.755		NS		

Note: NS – Not Significant @ $p > 0.05$

The summary result of the analysis shown in Table 6 indicates that $t = -0.313$ and $p = 0.755$. Since the significant value (p) is greater than the set value for the study ($p > 0.05$), the hypothesis that there is no significant difference between the mean conceptual knowledge of male and female Upper Basic 1 students in Mathematics knowledge measured using two-tier diagnostic test, is not rejected. This implies that there is no significant difference between the male and female mean's conceptual knowledge upper basic 1 students as assessed using two-tier diagnostic test. The conclusion drawn is that there is no gender bias in students' understanding of the mathematics concepts as measured using two tier diagnostic test in upper basic 1

5. Discussion of Findings

Findings arrived at in this research are discussed in this section.

5.1 Reliability and Validity of the Two-Tier Mathematics Diagnostic Test (TTMDT)

With the Rasch software program the researcher analysis of the TTMDT instrument of Upper Basic students' conceptual understanding and misconceptions of mathematics knowledge was done. The results from Table 1 showed that the item separation index (reliability) of upper basic one, two

and three students were 0.76, 0.74 and 0.68 respectively. These values are considered reasonable for research purpose. This finding is in conformity with the finding of Aliyu (2015) that person separation index of at least 0.50 is required for a separation of 1. Result in Tables 1, 2 and 3 further revealed that the items summary as the infit ranging from 0.9 – 1.3 and outfit 0.79 – 1.3. Also, the difficulty level ranges from -0.02 to 1 and 1.5 for part a and 1.5 to 2.5 for part b indicating that the items are not related, and measured different aspect of the mathematics knowledge. This finding corroborates the finding of Onah (2019) that showed low value of correlation coefficient as an indicator of test construct validity. The standard error of measurement (SEM) associated with the b-parameter of each of the Two-tier Mathematics Diagnostic Test (TTMDT) item is used to estimate its reliability. All items had SE within the range of 0.1 and 0.3. Thus, it can be concluded that the TEST was adequate in measuring the TTMDT. This finding is in agreement with the studies conducted by Lengkong, Istiyono., Rampean, Rejeki, Tumanggor and Nirmala (2020), Ramli, Talib, Hassan and Manaf (2020) Kanwal, and Farooq (2021) and Aliyu (2015) that two-tier diagnostic test is valid and reliable to distinguish students who understand the concepts from those who misunderstood, and with the use of Rasch analysis two-tier test items could be improved upon by rewording items that showed anomalies base on the individual-item fit.

5.2 Students' Conceptual Understanding of Mathematics

The finding from the result in Table 4 showed that students have few right conceptions in all the 20 items of the mathematics concepts under investigation. Only 1 student of Upper Basic 1, 4 students of Upper Basic 2 and none of Upper Basic 3 students respectively had at least 40 percent (%) of the items answered correctly, (That is, these number of students, were found to hold right conception. answering correctly both tiers: content and reason parts). Furthermore, result in Table 5 showed that for all the items, 58.62% of the students responded correctly at the tier 1 level, whereas only 24.21% responded correctly at both tiers indicating that on average, 24.21% of students had understanding of the mathematics concepts. 34.41% who scored correct responses at the tier 1 level, either guessed the answer or only had partial knowledge that they used to achieve the correct responses while 17.17% students lack the knowledge of the mathematics concepts. Also, result in Table 5 revealed that percentage of students who answered correctly the content part was higher than the percentage of students who answered both tiers of the test. The mean difference between

students' conceptual understanding (answering both tiers correctly) and their misconception (answering only content part) is 1.34 shown in Table 7 was significant as was confirmed in Table 15, indicating that students least understood the Mathematics concepts or lack understanding or have partial content knowledge. This may probably be because students may have memorized certain facts in mathematics knowledge without conceptual understanding (without knowing the how and why) This finding is in agreement with Onah (2019), Idehen and Omoifo (2015) and Likando and Ngoepe (2014) that the number of students holding correct conceptions of mathematics were fewer than the number holding misconceptions and the level of conceptual understanding of students in mathematics is weak which they attributed to their prior experience, preconceptions, cultural background, lower level or inadequate thinking and reasoning abilities which current teaching and assessment practices do not promote.

5.3 Gender Effects on Students' Level of Conceptual Understanding of Mathematics Knowledge

The result of the analysis in Table 6 showed that male and female students had a mean conceptual understanding score gain difference of 0.34. table 7, 8 and 9 also shows the comparison between male and female upper basic 1 with mean difference of 0.18, male and female upper basic 2 with mean difference of 0.55 and male and female upper basic 3 with mean difference of 0.21. However, these figures are considered small and confirmed in Table 16, 17 and 18 that the difference is not statistically significant with regard to mathematics knowledge under study. This implies that male and female students hold equal knowledge level as measured using two-tier diagnostic test. This finding can be related to the finding of Likando and Ngoepe (2014) who found that male and female students exhibit similar spatial ability in mathematics and hence performed similarly. This similarity in male and female students' understanding of mathematics knowledge may be as a result of their ability to mentally represent and manipulate objects in space and these skills predict better understanding.

6. Conclusion

Based on the findings of this study, the researcher concludes that two-tier diagnostic test indicate that this test could be used reliably and in a valid way to access students' conceptual understanding and identify the percentage of misconceptions of mathematics knowledge among upper basic students in Kogi state.

Furthermore, the use of two-tier mathematics diagnostic test (TTMDT) at any giving time of instruction helps a competent teacher to complete work perfectly in the class room. This is because when the students are aware that the underlying reasons for any answer is as important as the answer itself, more attention will be placed on it by the students themselves thereby promoting conceptual understanding of mathematics knowledge and misconception of mathematics knowledge will be eliminated. Also, students held more misconceptions in mathematics knowledge than their conceptual understanding signifying that students performed better when only first part of items of two-tier test (content part) was considered than when both parts of the items (content and reason parts) were considered. This is an indication that students have only partial knowledge of Mathematics knowledge without underlying reasoning. Gender was not significant in upper basic students' conceptual understanding of mathematics knowledge as assessed on TTMDT.

7. Recommendations

This study has shown that TTMDT is reliable and valid and can be used to assess students conceptual understanding and identified the percentages of misconceptions different from what is commonly used in achievement or teacher-made test which measured correct answer (content part) only. Based on these findings, the researcher recommends that:

Packages like Rasch model package, BILOG MG-3 and other similar packages should be made available to examination bodies and schools to analysis make their research works easily in the university libraries.

Teachers of mathematics should note the implication of misconceptions of mathematics knowledge and be encouraged to develop and use two-tier mathematics diagnose test in the process of teaching mathematics. This will help both the teacher and the students to identify such misconception and easy treatment.

The State Universal Basic Education Board (SUBEB) and Universal Basic Education board UBEC should initiate or support workshops, seminars and in-service training programmes aimed at giving serving teachers the opportunity to learn how to develop, interpret and use two-tier diagnostic tests these classes (schools) in helping them to use this type of testing to inform and improve mathematics teaching, thus learning.

The development and use of two-tier diagnostic tests should be incorporated in the mathematics and other subject curriculum for pre-service teachers. This will

enable the teachers in training institutions prepare for challenges ahead of their teaching profession.

References

- Akor, D.C., & Onah, F.O. (2015). Effects of laboratory approach on students' achievement in geometry in secondary schools in Education Zone C of Benue State Nigeria. *Benue Journal of Mathematics and Mathematics Education*, 1(4), 25-32.
- Alex, J. K., & Mammen, K. J. (2014). Gender difference amongst South African Senior Secondary School Learners' thinking levels. *Mediterranean Journal of Social Science*, 5(20), 1908-1914.
- Aligba, S. O., & Iorja, N. R. (2021). Assessment of conceptual and procedural knowledge of students with special needs in Mathematics in Benue State. *European Journal of Training and Development Studies*, 8 (1), 6-18.
- Aliyu, R.T. (2015) Development and validation of Mathematics achievement test using the Rasch model. An un published Thesis Submitted to the Postgraduate School in Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy (Ph.D) Degree in Measurement and Evaluation of the Delta State University, Abraka.
- Aljaghshi, A., S., Abu., S.F., Algutaibi, A.Y., & Mundt T. (2021). Validity and reliability of new instruments for measuring patient satisfaction with removable dentures. *National Library of Medicine*, 21(1), 446: doi:10.1186/51290321.
- Anjum, S. (2016). Gender differences in mathematics achievement and its relation with reading comprehension of children in upper primary stage. *Journal of Education and Practice*. 6, 71-75.
- Awofala, A.O.A. (2017). Accessing Senior Secondary School Students' Mathematical Proficiency as related to gender and performance in Nigeria. *International Journal of Research in Education and Science*, 3(2), 488-502. Doi: 10.21890/ijres.327908
- Ben-Hur, M. (2016). Concept Rich Mathematics Instruction. North Beauregard, St. Alexandria, VA. Retrieved December 28, 2016 from <https://www.ascd.org/pub.books/1703>
- Gafoor, K. A., & kurukkan, A. (2015). Why High School Students feel Mathematics Difficult? An Exportation of Affective Beliefs. DoI:10.13140/RG. 2.2.18880.12800

- Ghaicha, A. (2016). Theoretical framework for educational assessment: A Synoptic Review. *Journal of Education and Practice* online paper, 7(24).
- Gurel, D. K., Eryilmaz, A. & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconception in science. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 989-1008. Doi: 10.1297/Eurasia
- Idehen, F.O., & Omoifo, C.N. (2015). Development and validation of a two-tier multiple-choice diagnostic instrument for assessing senior secondary school students' conceptions of selected mathematics concepts. *International Journal of Research and Development (IJRD)*, 2 (1), 254-272.
- Ihekwaba, C., Nkwocha, P., C., & Unamba, E.C. (2019). Development and validation of students' attitude assessment scale for evaluating science process skills in Mathematics. *Journal of Evaluation*, 4(1), 100 – 112.
- Kanli, U. (2014). A study on identifying the misconceptions of pre-service and in-service teachers about basic astronomy concepts. *Eurasia Journal of mathematics, Science and Technology Education*, 10(5), 463-471., Doi: <http://dx.doi.org/10.12973>
- Kanli, U. (2015). Using a two-tier test to analyse students' and teachers' alternative concepts in astronomy. *Science Education International*, 26(2), 148 – 165.
- Kanwal, S., & Farooq, M. S. (2021). Development of diagnostic assessment test for misconceptions in mathematics at elementary school level. *Global Educational Studies Review*, <https://doi.org/10.31703/gesr> 4(1), 94-102.
- Kirbulut, Z.D., & Geban, O. (2014). Using three tier diagnostic tests to assess students' misconception of states of matter. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(5), 509-521.
- Lengkong, M., Istiyono, E., Rampean, B. A. O., Rejeki, A. M., Tumanggor., & Nirmala, M. F. T (2020). Development of two-tier test instruments to detect student's physics misconception. *Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences Advances in Social Science, Education and Humanities Research*: 528.
- Likando, K.M., & Ngoepe, M.G. (2014). Investigating Mathematics trainee teachers' conceptions of proof writing in algebra at Copperbelt College of Education, Zambia. *Mediterranean Journal of Social Sciences*, 5(14), 331-335.
- Malatjie, F & Machaba, F. (2019). Exploring mathematics learners' conceptual understanding of coordinates and transformation geometry through concept mapping. *EURASIA Journal of Mathematics, Science and Technology Education*, 15(12).
- Mastan, Z. P., Ahmad, N., N. & Yahya, F.A.N (2023). Factors influencing learners' engagement in the classroom. *Asian Journal of Research in Education and Social Sciences*, 2(5), 53-62.
- Mckenna, P. (2019). Multiple choice questions answering correctly and knowing the answering. *Interactive Technology and Smart Education*, 16(1), 59 – 73.
- Mehmood, K, Khalil, S.Z., Gul, N., Iqbal, A., & Kousar, N. (2023). Comparison of level of misconceptions in male and female students in mathematics and their possible reasons at primary level. *Journal of Positive School Psychology*, 7(6), 836-846.
- Merriam Webster Dictionary (2012). Definition of geometry. Retrieved 21st April, 2017 from <https://www.merriam-webster.com>
- Metha, G., & Mokhasi, V. (2014). Item analysis of Multiple-choice questions: An assessment of the assessment tool. *International Journal of Health Science and Research*, 4(7), 197-202
- Mill, S. (2016). Conceptual understanding: A concept analysis. *The Qualitative Report*, 21(3), 546-557. Retrieved 25th May, 2017 from <http://nsuwors.nova.edu/tqr/vol21/iss3/8>
- Onah, J. (2019), Measures of students' conceptual understanding and reasoning of Junior Secondary School in Geology using two-tier diagnostic test in Benue State Nigeria. An unpublished Doctorate Thesis, submitted to the Postgraduate School, Benue State University, Makurdi, Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy (Ph.D) in Mathematics Education.
- Ramli, N.F., Talib, O. Hassan, S.A., & Manaf, U.K.A (2020). Development and validation of an instrument to measure STEM teachers' institutional preparedness. *Asian Journal of University Education (AJUE)*, 16(3) <http://doi.org/10.24191/ajue>.
- Schademan, A. (2021). Promoting Student Conceptual Understanding of Mathematics in Elementary Classrooms Madison Frankina

- Liberal Studies Program, California State University, Chico EDTE 490 W: Liberal Studies Capstone
- Schartz, J. E. (2014). *Why do people have difficulty with mathematics?* City: Allyn Prentice Hall
- Thomas, L. (2023). simple Random sampling/Definition steps and Examples Scribbr. Retrieved October 4,2023 from <https://www.scribbr.com/methodology/simplerandom-sampling/>
- Timayi, J. M., Ibrahim, M. O., & Sirajo, A. M. (2016). Gender differentials in students' interest and academic achievement in geometry using jigsaw iv cooperative learning strategies. *Abacus: Journal of Mathematical Association of Nigeria*, 41(1), 147-157.
- Widiyatmoko, A., & Shimizu, k. (2018). The development of two – tier multiple choice test to assess students conceptual understanding about light and optical instruments. *Journal Pendidikan*,7(4), Doi:10.15294/71416591.
- Yang, X., Leung, F., & Zhang, S (2019) Junior Secondary School Students' Conceptions of and Approaches to Learning Mathematics and Their Relationships in Mainland China School of Mathematics and Statistics, Southwest University, Sustainability.